日英機械翻訳の改善に向けた日本語制限規則の構築と評価

宮田 玲† 影浦 峡† Anthony Hartley†‡
†東京大学教育学部／大学院教育学研究科
‡豊橋技術科学大学

1 はじめに

近年，コンピュータの進歩や電子化された言語資源の活用によって，機械翻訳の精度は徐々に改善されてはいる．しかし，とりわけ言語構造の大きく異なる日英間の機械翻訳においては，精度が十分とは言えない．そうした状況に対して，文書執筆プロセスの上流工程における原言語テキストの統制の必要性が指摘され，制限言語 (controlled language) や前編集 (pre-editing) など，翻訳精度の向上を試みる研究が進められている [1] [2].

本稿では日英方向の機械翻訳に着目して，仮説的に20種類の日本語制限規則を構築し，人手評価によりその効果を検証した．

2 関連研究

機械翻訳における日本語の制限規則に関しては，1980年代より長尾らのMAL(Machine Readable Language) [3] や吉田らの制限日本語 [4] などの先駆的な試みがある．長尾らは,「係り受け関係の制限」と「付属語及び連用中止法の用法の制限」により，文の曖昧さを可能な限り取り除くことを提案している．一方吉田らは，規格化文法の要件として,「表現したい内容が正確に記述できること」「文の構造・意味が一意的に定まる文法であること」「人間が読んで分り易いこと」「書き手に大きな負担がかからないで文書が作成できること」の4点を挙げている．

近年では，翻訳業界においてシンプリファイドテクニカルジャパニーズ [5] という日本語への制限規則が提案されている．小倉らは英文の品質が低いサンプルには「長い，複雑」「あいまい」「日本語固有の表現で書かれている」の傾向があると分析している．その上で英訳品質に直結すると想定される文法的要素をもとに仮説的にルールを開発し，それを検証することで，以下の6つのパターンを規制するルールセットを完成させた．

1. 動詞＋「ように」の組み合わせ
2. 「など」で例示を終える
3. 他動詞に対応する目的語があいまいである
4. 助詞「は」格が1文中に2回以上出現する
5. 副詞節が1文中に2回以上出現する
6. 自動詞に対応する主語がない

しかし，パターンが網羅されているとは言いがたく，客観的な評価も十分に提示されていない．機械翻訳をシステム全体として向上させるためには，原言語制限の要件の解明と適切な評価手法が求められる．特に先行研究では,「目標言語の品質」の観点からのみ制限規則が評価されていることが多いが，本研究では加えて「原言語の品質」の観点からも評価を試みる．

3 制限規則の構築

まず電子機器マニュアルの日英対訳コーパス[1]の日本語文を日英機械翻訳にかけ，その結果を人力による対訳英文と対比させながら，翻訳精度に関わると考えられる20種類の文法的特徴（F1–F20とする[2]）を抽出した．ここでは，誤字・脱字等の文法的に不適格な特徴については対象としなかった．また，F2, 6, 7, 12については，上記の小倉らの規定したパターンを参考にした．ただしこれら20種類の特徴で，翻訳精度に関わると想定される文法事項を網羅できたわけではなく，あくまで本マニュアルの特徴に依存した形での抽出である．続いて，各特徴を制限する形で，仮説的に20種類の制限規則を構築し，実際にセンテンスに適用するための前編集規則を作成した（表1）．

[1]「ヤマハ (株) デジタル楽器事業部商品開発部ユーザーリレーション G」作成・提供．1999-2011年度作成．このうち合計38527対の日英対訳文を調査対象とした．

[2]ただしF1については，文法的特徴というよりは，文章の構造的特徴である．

No	制限規則	前編集規則	編集前・後の日本語文サンプル
F1	1 文を 50 文字以内にする	分割する	[前] もともと、チャンネルごとに個別のセンド量を設定することができるわけですから、全体としては同じエフェクトを使っていても、それぞれ異なるエフェクトのかかり具合を実現することができるわけです。 [後] もともと、チャンネルごとに個別のセンド量を設定することができます。そのため、全体としては同じエフェクトを使っていても、それぞれ異なるエフェクトのかかり具合を実現することができるわけです。
F2	1 文が 3 つ以上の節をもたないようにする	分割する	[前] 電波が弱いときは、楽器をアクセスポイントの近くに移動し、電波が届く状態にしてください。 [後] 電波が弱いときは、楽器をアクセスポイントの近くに移動してください。そして、電波が届く状態にしてください。
F3	否定表現をなるべく使わない	肯定表現に変更する	[前] デスティネーションで設定したパラメーターがエレメントに関するものでないときは設定できません。 [後] デスティネーションで設定したパラメーターがエレメントに関するものであれば、設定できます。
F4	動詞+形式名詞「こと」を使わない	形式名詞「こと」を削除する	[前] さまざまな音量のフッテージを組み合わせて、独特なオルガンサウンドを作り出すことができます。 [後] さまざまな音量のフッテージを組み合わせて、独特なオルガンサウンドを作り出せます。
F5	形式名詞「もの」を使わない	a) 形式名詞「もの」を削除する b) 具体的な名詞に置き換える	[前] 接続コードおよび接続プラグは抵抗のないものをお使いください。 [後] 抵抗のない接続コードおよび接続プラグをお使いください。
F6	動詞+「ように」を使わない	「ように」を削除する	[前] このボタンをオンにすることにより、マスターエフェクトを本体サウンドにかけるようにします。 [後] このボタンをオンにすることにより、マスターエフェクトを本体サウンドにかけます。
F7	主題を表す副助詞「は」を使わない	他の助詞に置き換える	[前] ユーザーソングは 5 曲までしか記憶できません。 [後]5 曲までしかユーザーソングを記憶できません。
F8	曖昧な等位接続詞を避ける	読点を加える	[前] ファイルの読み込み元のカードまたはディスクを選択します。 [後] ファイルの読み込み元の、カードまたはディスクを選択します。
F9	助動詞「れる・られる」は可能用法で使わない	「できる」に置き換える	[前]1 つのウェーブフォームに割り当てられるキーバンクは最大 128 個までです。 [後]1 つのウェーブフォームに割り当てることができるキーバンクは最大 128 個までです。
F10	動詞「見える」を使わない	a)「見ることが (の) できる」に置き換える b) 別の動詞に変える	[前] ジョイントコネクターはプラスの面が外から見える方向に差し込みます。 [後] ジョイントコネクターはプラスの面が外から見ることのできる方向に差し込みます。
F11	複合名詞を避ける	助詞を補完する	[前] マーカー挿入ダイアログが表示されます。 [後] マーカーを挿入するダイアログが表示されます。
F12	副助詞「など」で例示を終えない	a) 副助詞「など」を削除する b) 副助詞「など」の後に名詞を補完する	[前] 市販のオーディオソングなどを USB 記憶装置に保存したものです。 [後] 市販のオーディオソングを USB 記憶装置に保存したものです。
F13	単独で接続助詞「たり」を使わない	対になるように接続助詞「たり」を補完する	[前] たとえば、音色を変更して違った雰囲気にしたり、適切なテンポに調節できます。 [後] たとえば、音色を変更して違った雰囲気にしたり、適切なテンポに調節したりできます。
F14	カタカナ語動詞を使わない	和語に置き換える	[前] 大切なデータはロードする前にフロッピーディスクにセーブしておきましょう。 [後] 大切なデータは読み込む前にフロッピーディスクに保存しておきましょう。
F15	接尾辞「感」を使わない	a) 接尾辞「感」を削除する b)「の感覚」に置き換える	[前] 音質とビート感のバランスを重視した設定です。 [後] 音質とビートの感覚のバランスを重視した設定です。
F16	動詞「かかる」を使わない	具体的な動詞に置き換える	[前] ハーモニー音にかかるビブラートの深さを設定します。 [後] ハーモニー音に適用されるビブラートの深さを設定します。
F17	動詞「なる」を使わない	a) 動詞「なる」を削除する b) 具体的な動詞に置き換える	[前] 鍵盤を弾いてから最大レベルになるまでの時間を設定します。 [後] 鍵盤を弾いてから最大レベルに達するまでの時間を設定します。
F18	動詞「行なう」を使わない	動詞「行なう」を削除し、名詞をサ変動詞化する	[前] オーディオ (ウェーブデータ) 録音を行なう場合はこちらをご参照ください。 [後] オーディオ (ウェーブデータ) 録音する場合はこちらをご参照ください。
F19	格助詞「で」を避ける	a) 他の助詞に置き換える b) 格助詞「で」を動詞に置き換える	[前] タブ切替 [E][F] ボタンで、再生したいソングが入っている場所を選びます。 [後] タブ切替 [E][F] ボタンを押し、再生したいソングが入っている場所を選びます。
F20	動詞「あります (ある)」を避ける	a) 動詞「あります (ある)」を削除する b) 具体的な動詞に置き換える	[前] コードの詳細は 64 ページにあります。 [後] コードの詳細は 64 ページに記載されています。

表 1: 制限規則と前編集サンプル

4 評価実験

4.1 実験データ

マニュアルの日英対訳コーパスを「1 文の長さが 50 文字以上」「1 文の長さが 50 文字未満」で大きく二分し，それぞれデータセット A（総数 10026 文），B（総数 28501 文）とする．F1 についてはデータセット A から，F2-F20 についてはデータセット B から，それぞれの特徴を含む日本語文を 5 つずつ準備した[3]．

続いて，抽出した 100 文の日本語文 (JA1 とする) に対して筆者自らが制限規則を適用し，100 文の編集後日本語文を準備した（JA2 とする）．編集の際は,「原則として意味を変えない」という前提のもとで,「1 つの文に対し，1 つの規則を当てはめる」ようにした．

JA1，JA2 それぞれ 100 文ずつを Systran[4]，excite 翻訳，Google 翻訳の 3 種類の機械翻訳にかけた翻訳英文を，それぞれ SY1, EX1, GO1, SY2, EX2, GO2 とする．さらに，JA1 に対する人力による対訳英文をコーパスから 100 文抽出して評価用データに加えた．

以上より，日本語文を合計 200 文，英語文を合計 700 文準備した．

4.2 評価手法

翻訳英文評価手法　「understandability（理解容易性）」の観点から,「fully（十分）・mostly（大体）・partly（一部）・not at all（全く）」の 4 段階で，1 文ずつ人手評価を行った．評価者は，英語を母国語とし，かつ原言語の日本語を知らない英国の大学生に依頼した．英文 700 文を 100 文ずつ，7 グループに分け，各グループにつき 4 名の評価者を割り当てた（合計 28 名）．各評価者は，50 文を 2 セット（合計 100 文）評価した．また評価は，Web アンケートによって実施した．

日本語文評価手法　「読みやすさ」の観点から,「かなり読みやすい・少し読みやすい・ほとんど変わらない・少し読みにくい・かなり読みにくい」の 5 段階で人手評価を行った．評価者は，日本語を母国語とする大学生に依頼し，31 名から回答を得た．各評価者につき，日本語文 JA1，JA2 の 100 ペア（200 文）を全て評価してもらった．評価は，同じく Web アンケートによって実施した．

[3]ある特徴を含む文がコーパス中に 5 文以上ある場合は，その中からランダムで 5 つ選択した．

[4]Version.6 を使用．登録されていない単語のみ筆者らが辞書登録した．

4.3 実験結果

翻訳英文評価結果　understandability について「fully・mostly・partly・not at all」の 4 段階評定を，それぞれ 4，3，2，1 点と置き換えた．各文につき，4 名の評価値の中央値を最終的な評価値とした．

評価は，それぞれの制限規則ごとに以下 2 つの観点から実施した．

1. 編集前後で評価値が何%「向上・低下」したか（相対評価）
2. 評価値が 3 以上を「合格」とみなし，編集前後で「合格率」は向上したか（絶対評価）

1. により，相対的な品質向上（あるいは低下）の割合を測定する．ここでいう「向上」とは，ある機械翻訳について，編集前 JA1 と編集後 JA2 の翻訳文（例えば，SY1 と SY2）の評価値を比較した時に，後者の方が「向上」していることを指す．各制限規則について，編集前の評価値 15（つまり SY1, EX1, GO1 の値が 5 つずつ），編集後の評価値 15（つまり SY2, EX2, GO2 の値が 5 つずつ）で合計 30 の評価値がある．ここでは，編集前の 15 の値の内，編集後に「向上した値」「低下した値」の数をそれぞれ数え，母数 15 に占める割合 (%) を出した[5]．

2. は，評価値 3 つまり「mostly understandable（大体理解できる）」レベルにどれだけ達しているかを測定する絶対評価である．編集前後それぞれ 15 の値の内，評価値 3 以上の値の数を数え，母数 15 に占める割合 (%) を出した．

表 2 より，F2, 7, 18, 20 については,「向上」の割合の方が「低下」の割合よりも 40%以上多い．合格率をみても，編集後で 20%以上上がっている．つまり,「1 文に 3 つ以上の節 (F2)」「主題を表す助詞　「は」(F7)」「動詞「行なう」(F18)」「動詞「あります（ある）」(F20)」を制限することで，解析の複雑さや文の曖昧さが減少し，機械翻訳の精度が向上することが明らかになった．

一方，F3, 14 については合格率が 20%以上下がっており，また「低下」の割合の方が「向上」の割合よりも多い.「否定表現をなるべく使わない (F3)」という規則のもとで，半ば無理やりに肯定表現に書き換えたことが逆効果になったと言える．またカタカナ語動詞はむしろ英語に即した表現であるので，単純に「カタ

[5]ただし，F4 に関しては，原文に不備があったため，編集前後それぞれ 12 ずつの評価値がある．

カナ語動詞を使わない (F14)」とするのではなく，範囲を定めて規制する必要があるだろう[6].

日本語文評価結果　読みやすさについて，編集後の文の方が「かなり読みやすい・少し読みやすい・ほとんど変わらない・少し読みにくい・かなり読みにくい」の 5 段階評価を,「2，1，0，-1，-2」点に置き換えた.

評価は「編集前後でどの程度読みやすさが変化しているか」の観点から行った（相対評価）. JA1，JA2 の 100 ペアについて，それぞれ 31 の評価値がある．日本文の読みやすさの評価においては，統計的な有意差を重視して Wilcoxon 符号付順位和検定を行った．各制限規則 5 ペアについて，5%有意水準で，制限によって読みやすさが「向上したペア」と「低下したペア」の数を数え，それぞれ母数 5 で割った[7]．以下，日本語文の読みやすさに関して「向上」「低下」という場合は，5%有意水準での差があるものに限る.

表 2 より，F2, 3, 5, 7, 8, 10, 12, 15, 16, 17, 19, 20 については,「低下」割合が「向上」割合よりも大きい．特に F2,5,7,10,16,17 は，編集によって 60%以上の文の読みやすさが「低下」しているを示している．今回は「原則として意味を変えない」という前提で編集しているが，分割・語順変更・削除といった編集によって多くの日本語表現が不自然になったと考えられる．読みやすさの低下の度合いがどこまで許容可能であるかを絶対評価により調査することが今後必要だろう.

一方，F13,14 は「向上」の割合が「低下」の割合よりも大きい.「単独で使用される接続助詞「たり」(F13)」については,「〜たり，〜たり」の形の方が日本語母語話者にとって一般的な表現であり，また「カタカナ語動詞 (F14)」については，カタカナ語よりも和語の方が馴染みがあるためであろう.

5　おわりに

本研究では，機械翻訳しにくいと考えられる文法的特徴を 20 種類抽出し，それに対する制限規則・前編集規則を仮説的に構築した．また，実験により翻訳英文と日本語文の双方の品質の観点から評価を行った．いくつかの規則は機械翻訳の精度向上に効果があったが，よりサンプルを増やして検証することが必要である．今回効果が出なかった規則についても，個別に診断して改善を試みる予定である．また，全体的に日本語の読みやすさが低下する結果となったため，今後は翻訳精度を向上させるとともに，日本語の読みやすさをなるべく低下させないような前編集規則を構築することが必要であろう．制限規則の拡張と同時に，各規則の精緻化を進めることが今後の課題である.

[6]例えば「アピールする」など英語の語義と異なる表現
[7]F4 に関しては，母数 4

No.	翻訳英文 変化率 (%) 向上	翻訳英文 変化率 (%) 低下	翻訳英文 合格率 (%) 編集前	翻訳英文 合格率 (%) 編集後	日本語文 変化率 (%) 向上	日本語文 変化率 (%) 低下
F1	47	33	13	7	20	20
F2	67	20	27	47	0	100
F3	33	47	40	20	20	40
F4	25	42	50	50	25	25
F5	40	40	27	53	0	60
F6	33	20	20	20	20	20
F7	60	20	47	67	20	80
F8	53	40	47	33	0	40
F9	47	27	20	33	0	0
F10	40	40	33	27	40	60
F11	40	53	40	27	20	20
F12	40	27	47	47	0	20
F13	33	27	13	20	60	0
F14	20	40	53	33	60	20
F15	13	47	7	7	0	40
F16	40	27	47	40	0	80
F17	27	47	40	33	20	60
F18	67	20	20	47	20	20
F19	33	33	40	27	0	40
F20	60	13	27	73	20	40

表 2: 翻訳英文・日本語文の評価結果

謝辞　研究用の対訳データは，YAMAGATA INTECH 株式会社の中村哲三様，ヤマハ株式会社の遠藤幸夫様にご提供いただいた．また本研究の一部は，日本学術振興会科学研究費補助金基盤 (A)「包括的な翻訳情報資源を実現する統合翻訳支援サイトの構築」(課題番号 00211152）の支援を得て行われた.

参考文献

[1] H. Kaji. “Controlled Languages for Machine Translation: State of the Art,” *Proceedings of MT Summit VII*, pp.37–39(1999).

[2] A. Bernth, C. Gdaniec. “MTranslatability,” *Machine Translation,* vol.16, pp.175–218(2001).

[3] 長尾真, 田中伸佳, 辻井潤一 “制限文法にもとづく文章作成援助システム” 『情報処理学会研究報告（NL)』vol.1984, no.27, pp.1–8(1984).

[4] 吉田将, 松山晶子 “日本語の規格化—係り受け関係の規格化とそれへの変換ルール—” 『情報処理学会研究報告（NL）』vol.1985, no.31, pp.1–6(1985).

[5] 小倉英里, 工藤真代, 柳英夫 “シンプリファイド・テクニカル・ジャパニーズ英訳を視野に入れて日本語を作る” 『情報処理学会研究報告（DD)』, vol.2010, no.5, pp.1–8(2010).